The effect of sociodemographic (mis)match between interviewers and respondents on unit and item nonresponse in Belgium

Anina Vercruyssen [a]

Celine Wuyts [a]

Geert Loosveldt [a, b]

ABSTRACT

Interviewer characteristics affect nonresponse and measurement errors in face-to-face surveys. Some studies have shown that mismatched sociodemographic characteristics – for example gender – affect people's behaviour when interacting with an interviewer at the door and during the survey interview, resulting in more nonresponse. We investigate the effect of sociodemographic (mis)matching on nonresponse in two successive rounds of the European Social Survey in Belgium. As such, we replicate the analyses of the effect of (mis)matching gender and age on unit nonresponse on the one hand, and of gender, age and education level (mis)matching on item nonresponse on the other hand. Recurring effects of sociodemographic (mis)match are found for both unit and item nonresponse.

Keywords: interviewer characteristics, interviewer effects, unit nonresponse, item nonresponse, sociodemographic matching

a) Centre for Sociological Research, KU Leuven - University of Leuven

b) Corresponding author:
Prof. Dr. Geert Loosveldt
Centre for Sociological Research, University of Leuven
Parkstraat 45 - bus 3601
3000 Leuven
Belgium
tel. *+32 16 32 31 54*
fax *+32 16 3 23365*
e-mail to: Geert.Loosveldt@Soc.KULeuven.be

The effect of sociodemographic (mis)match between interviewers and respondents on unit and item nonresponse in Belgium

Abstract

Interviewer characteristics affect nonresponse and measurement errors in face-to-face surveys. Some studies have shown that mismatched sociodemographic characteristics – for example gender – affect people's behaviour when interacting with an interviewer at the door and during the survey interview, resulting in more nonresponse. We investigate the effect of sociodemographic (mis)matching on nonresponse in two successive rounds of the European Social Survey in Belgium. As such, we replicate the analyses of the effect of (mis)matching gender and age on unit nonresponse on the one hand, and of gender, age and education level (mis)matching on item nonresponse on the other hand. Recurring effects of sociodemographic (mis)match are found for both unit and item nonresponse.

1. Introduction

It is widely known that interviewer characteristics can influence survey data collection processes. Recent studies also provide evidence that the sociodemographic characteristics of interviewers can affect the likelihood of co-operation (e.g. Durrant et al. 2010) and the quality of respondents' answers (e.g. Oyinlade and Losen 2014; Samples et al. 2014). However, the effect on survey errors of matching the sociodemographic characteristics of interviewers with those of sample units in face-to-face surveys still requires further study. Some scholars have pointed out that we need a greater variety of sociodemographic characteristics and outcome

variables in order to understand the effects of a (mis)match between the interviewer and sample unit, and the possible resulting survey errors (Davis et al. 2010). In line with this recommendation, in the current study we investigate the effects of (mis)matching interviewers and sample units with regard to gender, age and education level on several nonresponse indicators in the Belgian data of the sixth and seventh round of the European Social Survey (ESS6BE and ESS7BE).

2. Effects of sociodemographic (mis)matching on unit nonresponse

The idea of matching the characteristics of interviewers and sample units in face-to-face surveys is based on the theory of liking (Groves et al. 1992) and the concept of social distance (Williams 1964). The theory of liking states that people tend to prefer interaction with people who they like or with whom they share similarities. The concept of social distance is also an expression of (dis)similarity between individuals. It has been operationalised as belonging to different social classes (e.g. Katz 1942), having different ethnic backgrounds (e.g. Weeks and Moore 1981), or a combination of race and social rank or class (e.g. Williams 1964). Not only race and social status may serve as a basis of perceived similarities between people; the tendency to prefer interaction with similar people is also influenced by other sociodemographic characteristics, such as gender and age (Lazarsfeld and Merton 1954; Lipman-Blumen 1976). Accordingly, we could expect that when interviewers and sample units are more alike in terms of sociodemographic characteristics, the sample units will act more cooperatively in response to a survey participation request and during a survey interview.

Regarding the effects of the sociodemographic characteristics of interviewers, regardless of the characteristics of the sample units, existing literature shows mixed results. Some studies

have found no significant effects of the gender, age or education level of the interviewer on unit nonresponse (e.g. Hox et al. 1991; Pickery and Loosveldt 2002, Pickery, Loosveldt & Carton 2001), whereas others have found effects, but not always exactly the same ones (e.g. Durrant and D'Arrigo 2014; Hox and de Leeuw 2002; Hansen 2007; Singer et al. 1983). Comparing studies on the effects of age is hampered by the variety of operationalisations, mostly binary distinctions with different cut-off points for comparing old versus young interviewers (Freeman and Butler 1976; Lipps 2010; Wilson et al. 2002). In addition to other factors, interviewer experience is also known to affect survey participation levels (e.g. Durrant et al. 2010; Hansen 2007).

When also considering the characteristics of the sample units, previous research on the effect of (mis)matching sociodemographic characteristics on survey nonresponse has predominantly focussed on (mis)matching gender and race/ethnicity (for an overview, see e.g. Davis et al. 2010; Oyinlade and Losen 2014). Some studies have found that racial mismatching can lead to lower co-operation rates (e.g. Moorman et al., 1999), whereas others have not found any significant differences (e.g. Singer et al. 1983). With regard to gender (mis)matching, some studies have shown that gender matching improves the likelihood of co-operation (Durrant et al. 2010) and decreases the likelihood of dropouts during the interview (Catania et al. 1996), whereas other research has found no significant effects (Campanelli and O'Muircheartaigh 1999). With regard to age, Lipps (2010) has found that older male interviewers (> 30 years of age) were significantly better at convincing older male as well as young female sample units to participate. Education (mis)match between the interviewer and the sample unit has not been studied to the same extent as gender and racial (mis)matching in the context of survey nonresponse. In a recent study, Durrant and colleagues (2010) did find that education mismatch can affect the likelihood of refusal in surveys. The negative effect of education

mismatch was observed both for the interviewer having attained a higher education level and for the respondent having attained a higher education level.

This brings us to our first research question: are sample units more likely to participate if the interviewer is more similar to them in terms of gender and age? A good operationalisation of matching is also crucial in this context.

3. Effects of sociodemographic (mis)matching on item nonresponse

The sociodemographic characteristics of interviewers not only affect unit nonresponse, but also influence how and what respondents answer during an interview (for an overview, see e.g. Hox et al. 1991; Davis et al. 2010). Of course, item nonresponse is also to a large extent affected by the characteristics of the respondents themselves with miscomprehension of the question, failure to retrieve adequate information and unwillingness to disclose information as the main reasons for item nonresponse (Groves et al., 2009; Tourangeau et al., 2000). It is also not hard to imagine that elderly people may more often have issues remembering things and that knowledge questions tend to be easier for those with a higher education level. There are plenty of studies investigating the effects of the characteristics of interviewers on the answering process in surveys, on the one hand, and for the effects of the characteristics of the respondent, on the other hand. Research on how (mis)matched sociodemographic characteristics of interviewers and respondents affects the latter's answers, however, is relatively less abundant than research on how it affects unit nonresponse. It is also sometimes dated. Durrant and colleagues (2010) suggested that such studies on item nonresponse errors may be limited because many surveys do not collect detailed information on the interviewers and/or employ only a small number of them.

Respondents' answers to factual questions can be influenced by the interviewer's sociodemographic characteristics (O'Muircheartaigh and Campanelli 1998), but subjective and sensitive questions tend to be even more vulnerable to these effects (see Davis et al. 2010; Schnell and Kreuter 2005). When also taking into account the sociodemographic characteristics of respondents, studies have mostly focussed on race and/or gender (for overviews, see e.g. Davis et al. 2010). A mismatch of race and/or gender has been found to affect the frequency of affirmative answers to (attitudinal) questions (Oyinlade 2014; Oyinlade and Losen 2014).

Moreover, sociodemographic (mis)match influences the tendency to want to disclose attitudes to interviewers in the first place. For example, women appear to be more open to female interviewers (Fletcher and Spencer 1984). With regard to more sensitive questions – for example those concerning psychological issues – women and men appear to be more open to female interviewers (Chun et al. 2011; Pollner 1998) and racial mismatch reduces the likelihood of reporting psychological issues (Samples et al. 2014). Mixed results have been found regarding the effect of gender matching on questions about sexual behavior (Catania et al. 1996; Wilson et al. 2002) and undesirable behavior such as drug abuse (Johnson and Parson 1994). Studies on the effects of age and education (mis)matching are harder to find than those on gender and racial matching. Age matching may make older people feel more at ease (Rodgers and Herzog 1992) and makes younger respondents report larger personal networks (van Tilburg 1998). With regard to sensitive questions about sexual behavior, mixed results are found again: some sexual behavior is reported more to older interviewers and other behavior more to younger ones (Wilson et al. 2002).

The effect of education (mis)match on item nonresponse in surveys has barely been studied. When it has been investigated, education level forms part of a composite measurement of socioeconomic status (SES) that combines an education level score with occupation-related indicators, such as the skill level and prestige of the occupation (van Tilburg 1998), which can mask the specific effects of education level itself. SES indicators have also been used in research into the effects of social distance between interviewers and respondents, but these studies have mostly focussed on the occupation of respondents and are dated (e.g. Katz 1942; Williams 1964), leaving the effects of education (mis)match between interviewers and respondents on data quality under-explored in survey methodological research.

Based on the 'theory of liking', we would expect that interviews proceed more smoothly when there are more sociodemographic similarities between the respondent and the interviewer, resulting in less item nonresponse. Hence, our second research question is whether educational (mis)match between interviewers and respondents, in addition to gender and age (mis)match, affects item nonresponse. We test the effect of (mis)match on item nonresponse using data from two rounds of the European Social Survey (ESS). Before investigating the effects on item nonresponse, we address our first research question regarding sociodemographic (mis)matching and unit nonresponse in the form of refusal.

4. Methods

4.1. Data

We test whether sociodemographic (mis)matching affects unit and item nonresponse using data from round 6 (ESS6BE) and round 7 (ESS7BE) of the European Social Survey (European Social Survey 2012, 2014) in Belgium. The standard ESS data release includes sample units' and interviewers' age and gender, allowing an assessment of the effects of age

and gender (mis)matching on both unit and item nonresponse. The ESS data further includes respondents' education level and can be linked to the Belgian interviewers' education level. Interviewers' education level is not part of the standard ESS data but is separately collected by the Belgian national team via an interviewer questionnaire at the start of the fieldwork. This data allows an assessment of the effects of education (mis)matching on item nonresponse among respondents, which have been rarely examined. The education level of non-respondents remains unfortunately unavailable, preventing a corresponding assessment of the effects of education (mis)matching on unit nonresponse.

The two rounds were implemented in accordance to the same design and specifications (see Tirry and Loosveldt 2013 and European Social Survey 2016a for more information on ESS6BE and Barbier et al. 2016 and European Social Survey 2016b for more information on ESS7BE). Sample units are randomly drawn from the Belgian National Register. Sets of geographically proximate sample units are assigned to interviewers mainly on the basis of geographic and temporal need and interviewer availability. Sociodemographic characteristics of interviewers and sample units are not taken into account in the assignment of sample units to interviewers. All interviewers can therefore be expected to encounter sample units with diverse sociodemographic backgrounds.

In order to encourage a sufficiently large number of these sample units to participate in the survey, the interviewers are required to adhere to a strict contact procedure. The first attempt to contact a sample unit should always be face-to-face, and at least five (face-to-face) contact attempts should be made, spread over different days of the week (at least one attempt during the weekend), over different times of the day (at least one attempt after 5PM), and over at least two different weeks. Non-respondent sample units that are not hard refusals may subsequently be reassigned to another interviewer, who is then required to repeat the entire

contact procedure. The same fieldwork agency was employed for both rounds and approximately 60% of the interviewers working in ESS6BE also worked in ESS7BE.

Taking advantage of the similarities in design and implementation, we conduct the analysis separately for both ESS rounds, as replicate samples drawn from a highly similar context, in order to assess the stability of the results.

4.2. Variables

Unit nonresponse

The main survey response outcome of interest in this study is refusal among sample units that were successfully contacted. In the context of ESS, a refusal by the sample unit himself or herself (not by proxy) almost always implies a (brief) face-to-face interaction between the interviewer and the sample unit, in which some sociodemographic characteristics could be observed. Interviewers are strongly discouraged to make contact attempts by telephone. We therefore compare the sample units that refused participation to the other eligible sample units that were successfully contacted (respondents and sample units categorized as 'unable or other nonresponse').

Yet refusal is only one of the components of unit nonresponse. Unit nonresponse can occur for various reasons other than the sample unit being reluctant to participate. Considering unit nonresponse in a broader sense, we compare the sample units that actually participated to all the other eligible sample units (respondents, sample units categorized as 'unable or other nonresponse' and noncontacted sample units). Although we do not expect strong effects on the other components of unit nonresponse, we cannot a priori assume they are non-existent. For example, the interaction between interviewers and sample units may affect the latter's perceived physical or mental ability, and their being categorized as 'not able'. And although by definition no interaction could have taken place between interviewers and noncontacted

sample units, interviewers are informed on the gender and age bracket of all sample units assigned to them, and may alter their effort and/or the timing of contact attempts accordingly.

Reassignment of non-respondent sample units to a second or even third interviewer complicates the unit nonresponse analysis. For re-assigned sample units, several interviewers may jointly have affected the final survey response outcome. We therefore restrict the analysis to the initially assigned interviewers and the survey response outcomes resulting from their efforts. The distribution of survey response outcomes of the initially assigned interviewer (i.e., before any reassignments) and the outcomes at the end of the fieldwork are presented in Table 1.

[Table 1 around here]

Item nonresponse

Item nonresponse in ESS6BE and ESS7BE results from respondents either refusing ('Refusal') to answer a particular question, or being unable to ('Don't know'). Table 2 shows the top ten items with the highest item nonresponse in ESS6BE and ESS7BE. Seven items in the ESS main questionnaire appear in both rounds. This shared set of items with high item nonresponse consists of household income, education level of both the parents and partner, occupation of the mother when the respondent was 14 years old, political party vote and political orientation on a left-right scale. The relatively high item nonresponse for these items observed in both ESS6BE and ESS7BE, especially compared to the low item nonresponse observed for most other items, suggests that these questions are consistently more prone to item nonresponse. Some of them are sensitive (e.g. political party vote), some are likely to cause retrieval or knowledge issues for the respondents (e.g. education level of the parents), and others, such as household income, may be both sensitive and difficult. We will focus on

this set of seven shared items for both surveys to investigate whether recurring patterns of effects of sociodemographic (mis)matching could exist.

Item nonresponse occurs for a vast majority of the interviewers in ESS Belgium at some point or other. In both rounds, about 80% of the interviewers registered a Don't know or Refusal for one or more of this set of items in at least one interview he or she administered (not tabulated). It can therefore not be specifically attributed to a very small number of interviewers consistently skipping particularly sensitive or difficult questions, or otherwise (unconsciously) encouraging respondents to abstain from answering particular questions.

As an overall measure of item nonresponse, we construct a binary indicator to compare respondents for whom none of these seven items is missing to respondents for whom at least one is missing. About one in four respondents in both ESS6BE and ESS7BE belongs to the latter group.

[Table 2 around here]

(Mis)matching of sociodemographic characteristics

A descriptive summary of gender, age and education level for the sample units (including both respondents and nonrespondents), the respondents and the interviewers is shown in Table 3. Matching variables were constructed for gender and age for all sample units with respect to the initially assigned interviewer, and for gender, age and education level for respondents with respect to the interviewer who administered the questionnaire. The frequency of sociodemographic (mis)matches is summarized in Table 4.

[Table 3 around here]

[Table 4 around here]

Gender matching between interviewers and sample units/respondents leads to four possible categories, with two matches (male interviewer - male sample unit, female interviewer - female sample unit) and two mismatches (male interviewer - female sample unit, female interviewer - male sample unit). In both ESS6BE and ESS7BE, about one in two sample units were initially assigned to an interviewer with the same gender (Table 4).

Age matching is determined on the basis of the age difference between interviewers and sample units/respondents. Rather than choosing a cut-off point to distinguish between 'young' and 'old' or between age cohorts, we define matching as a maximum age difference of five years in either direction. In both ESS6BE and ESS7BE about half of the sample units were initially assigned to an older interviewer and only about one in six was assigned to an interviewer with roughly the same age (Table 4).

Education level matching unfortunately cannot be determined for all sample units with respect to the assigned interviewers, as this information is retrieved from the survey and thus unavailable for non-respondent sample units. Education level matching can only be determined for the respondents with respect to the interviewers who administered the questionnaire. For the interviewers, four education levels are distinguished: less than higher-secondary education (secondary education not completed), (higher) secondary education, higher education at college level, and higher education at university level. We recoded the more detailed respondent education level variable that is included in the ESS main data to these four categories. As with age matching, we adopted a straightforward operationalisation into three matching categories: respondent and interviewer have the same education level (match), the respondent has a higher level, or the interviewer has a higher level. The predominant combination is interviewers with a higher education level than their respondents. For more than half the respondents the questionnaire was administered by an interviewer with a higher education level (Table 4).

4.3. Modelling approach

A two-level logistic random intercept model, with sample units nested within interviewers, is estimated for each of the two unit nonresponse indicators (unit nonresponse due to refusal among sample units that were successfully contacted and overall unit response). Likewise, a two-level logistic random intercept model, with respondents nested within interviewers, is estimated for the item nonresponse indicator (item nonresponse to at least one of the seven shared items with high item nonresponse). The analysis is repeated for ESS6BE and ESS7BE.

In addition to the sociodemographic matching variables as the main explanatory variables of interest, some additional characteristics of interviewers and sample units/respondents are included in the model estimation.

We take interviewer experience into account because older interviewers are likely to have more experience working as an interviewer and interviewer experience has been shown to affect nonresponse (e.g. Durrant et al. 2010; Hansen 2007). Interviewer experience is collected via the short interviewer questionnaire at the start of the fieldwork. It is included in the analysis as a categorical variable with four categories: less than one year of experience, 1 up to 5 years of experience, 5 up to 10 years of experience, and more than 10 years of experience. The distribution of interviewers' experience is included in Table 3.

Sample units' age and education level can affect whether they understand survey requests – and among the respondents, the survey questions (Bethlehem et al. 2011; Groves and Couper 1998). We therefore additionally control for these characteristics. Age is available for all sample units; education level is only available for respondents. The gender of all sample units is included in the gender match variable.

5. Results

We start by looking for an answer to our first research question: do sample units participate more if the interviewer is more similar to them with regard to gender and age? In Table 5, we find some support for the theory of liking interaction with similar people. In ESS6BE, male interviewers assigned to female sample units have significantly more refusals compared to situations where there is male gender matching. The probability of response is also lower for this group but the effect is not statistically significant at the traditional level of statistical significance of 5%. When we use the female-female matching as reference category, male interviewers assigned to female sample units also have a significantly lower response and more refusals in ESS6BE than when female interviewers are assigned to female sample units (not tabulated, respectively, b = -0.280, b= 0.320, both $p < .05$). This shows that gender mismatching can have disadvantageous effects on survey participation. For ESS7BE, the signs of the male interviewer-female sample unit coefficients suggest a similar pattern, but the gender matching effect is not statistically significant.

We also see a significant effect of age matching in ESS7BE. When the interviewers are older than the sample units, response is significantly lower than when the interviewers and sample units are of a similar age (+/- 5 years). Age similarity between interviewers and sample units may be beneficial for survey participation in some situations, but it does not significantly affect the probability of refusal or overall response in ESS6BE. The effects of education level (mis)matching could unfortunately not be tested as the information is not available for nonrespondents. However, we did include the education level and years of experience of the interviewers. Interviewers' education level is not significantly related to unit nonresponse in either round. The effects of interviewers' experience level are mixed. In ESS6BE, sample units assigned to more experienced interviewers tend to be less likely to participate. ESS6BE shows a reversed pattern with sample units assigned to more experienced interviewers being more likely to participate.

Although Table 5 shows interesting effects of sociodemographic matching on survey participation, and adding the matching variables significantly improves the fit at least of the refusal model in ESS6BE and the response model in ESS7BE, significant differences between interviewers for response and refusal remain unexplained by the models, as can be seen in the large intercept variances.

[Table 5 around here]

[Table 6 around here]

Next, we address the second research question regarding whether interviews have less item nonresponse when interviewers and respondents have matching sociodemographic characteristics. The results are shown in Table 6. We find significant effects of gender matching and age matching in both ESS6BE and ESS7BE. However, only male gender matching reduces item nonresponse whereas female interviewers paired with female respondents are significantly more likely to have item nonresponse in comparison to male gender matching. This only gives partial support for the 'theory of liking' with regard to gender: shared gender only produces less item nonresponse among men. More support for the 'theory of liking', however, can be found when we look at age matching. Having a similar age also significantly reduces item nonresponse. Given that we control for the age of the respondent, the significant age matching effect cannot purely be attributed to respondents being elderly people with retrieval issues. On the contrary, in ESS7BE we also see a direct effect of age with more item nonresponse when the interviewers are older. The age matching effect is also not due to the interviewers' (in)experience.

The data also allowed us to test the effects of education level matching, which has not often been tested yet in this context. With regard to the effects of education (mis)matching, we observed significant effects on item nonresponse, with interviewers having a higher education level having a positive effect and interviewers having a lower education level having a negative effect, when only the matching variables were included in the models (not tabulated). However, Table 6 with the models with the control variables shows that none of these effects are actually education-level (mis)matching effects. They are in fact effects due to the education level of the respondents: those respondents with the lowest education level have significantly more item nonresponse among the shared items. Interesting to note with regard to Table 6 is that highly similar results are found when we focus on the complete top ten of items with high nonresponse for both surveys separately (not tabulated).

Despite all the interesting, significant effects of sociodemographic matching on item nonresponse, and the significant improvement in model fit from adding the gender matching variables, in both rounds of ESS in Belgium, much of the variance at the interviewer level remains unexplained.

6. Discussion

Previous research has shown that sociodemographic (dis)similarities between interviewers and sample units can affect unit and item nonresponse in surveys. However, existing research remains inconclusive about the effects of (mis)matching. Our study adds to the body of literature by assessing the effect of mismatched sociodemographic characteristics between interviewers and sample units on both unit and item nonresponse in two highly comparable surveys and by using a straightforward operationalisation of sociodemographic (mis)matching. The availability of information about the education level of the interviewers

allowed us to test the effects of education mismatch between interviewers and respondents on item nonresponse, which has been infrequently investigated in previous research. We did find several effects of sociodemographic (mis)matching in the two ESSBE surveys. However, they are only consistent for item nonresponse, suggesting that some shared sociodemographic characteristics between interviewers and respondents may lead to face-to-face interviews that proceed more smoothly, as proposed by the theory of liking and the concept of social distance.

For our first research question on the effects of sociodemographic matching on unit nonresponse, we find that female gender matching has some beneficial effects with regard to preventing refusals in ESS6BE. This is in line with people preferring interaction with similar people, as in the findings of Durrant and colleagues (2011) on gender matching and survey co-operation. However, for ESS7BE we do not find significant effects of gender matching. For that survey, we do find beneficial effects of age matching on response in comparison with situations where the interviewer is older than the sample unit. This finding contrasts other studies' finding that older interviewers may radiate more authority and hence have fewer refusals (e.g. Blom et al. 2010). As the sociodemographic matching effects on unit nonresponse are not cumulative, testing an alternative operationalisation of sociodemographic distance with counting of the number of matching sociodemographic characteristics did not work particularly well. Such a count variable would also conceal the sometimes opposite or absent effects of the separate (mis)matching characteristics on unit nonresponse.

The findings do suggest that it is relevant to continue monitoring the effects of gender and age (mis)match on survey response. Further research is needed to determine whether survey practitioners should consider matching interviewers and sample units in face-to-face surveys or, at least in the refusal-conversion phase. The observation that female sample units are more likely to refuse participation when assigned to a male interviewer as opposed to a female

interviewer in ESS6BE appears intuitive: Women may be less inclined to let an unknown man into their home, and refusal conversion could take that into account.

The effect of education matching on unit nonresponse could not be tested, as we unfortunately do not have data for the education level of nonrespondents. The education level of the interviewers in itself did not affect response and refusal. With regard to the possible effects of education level (mis)matching on unit nonresponse, sample units can, of course, not estimate an interviewer's education level as straightforwardly as they can gender or age. Nevertheless, in the (brief) face-to-face interaction between interviewers and refusing sample units, the sample units may very well be able to make (unconscious) judgements about whether their own level of education corresponds with that of the interviewer, based on for example language use. Future studies should consider this.

Regarding the interactions with the interviewers, we also need to point out that our study specifically focussed on face-to-face interviews. In telephone interviews, sample units and respondents only hear the voice of the interviewer, making it harder for them to derive the sociodemographic characteristics of the interviewer and potential similarities with their own characteristics. Future research should investigate whether and how sociodemographic similarities affect the interactions in telephone surveys and whether and how such effects differ to the effects in face-to-face surveys.

In answer to our second research question concerning sociodemographic matching and item nonresponse, we find stable effects of gender and age matching in both surveys. Male gender matching appears to make interviews proceed more smoothly and to produce less item nonresponse, whereas female gender matching appears to increase item nonresponse in both surveys. These significant gender-matching effects on item nonresponse also remain stable in both surveys after the control variables are added. For both surveys we also see less item nonresponse in the case of age matching than when interviewers are younger than their

respondents. In contrast to expectations, matching education levels does not appear to relate to item nonresponse.

With regard to the ambiguous effect of gender matching on item nonresponse, there may be an explanation that is in line with the theory of liking as well as with studies finding that women are more at ease with female interviewers (e.g. Chun et al. 2011; Pollner 1998). A possible explanation is that female interviewers may show more understanding when a female respondent gives them clues that she does not wish to answer a question. Future research should investigate whether women may sometimes become too comfortable together in interview situations, with female interviewers refraining from probing to avoid female respondents becoming uncomfortable. If this proves to be the case, special attention could be paid to this in interviewer training. The results of this study also indicate that solely distinguishing between gender matching or not, would obscure the further differences between male-male and female-female situations.

With regard to the prevention of item nonresponse, our study also provides interesting insights in addition to the effects of sociodemographic (mis)matching. First, we see that there are not that many questions with extreme amounts of item nonresponse, which is a positive finding concerning the data quality. A second observation is that out of the top ten nonresponse items, seven occur in both of the surveys. Possibly the interviewer training could pay more attention to these specific questions, as the recurring problem of item nonresponse for these seven items does not seem to be a chance occurrence. Third, item nonresponse does not seem to be caused by the skills of a handful of (unsuccessful) interviewers in ESSBE, as can be seen in the number of interviewers confronted with item nonresponse in the top ten questions. About 60% of the interviewers worked in both ESS6BE and ESS7BE, but the effects of (mis)matching on item nonresponse are spread between this group and the interviewers who only worked in one of the two surveys, as became apparent when we

investigate the groups separately (not tabulated). Accordingly, item nonresponse seems not to be specifically related the interviewers who worked on both surveys.

Sociodemographic matching seems to be a component in the explanation of unit and item nonresponse in Round 6 and Round 7 of ESS in Belgium, although more consistently so for item nonresponse. Still, significant interviewer variance remains in all models, indicating that gender matching constitutes only a small piece of the nonresponse jigsaw. We also hope to inspire more research into the topic, to learn whether similar effects can be found in other countries. On a more practical level, we hope it will lead to more investigation into how to optimize the allocation of sample units to interviewers, not only for ESS but for face-to-face surveys in general as well.

Acknowledgements

This study was funded by the Hercules Foundation – Research Foundation-Flanders (FWO).

References

American Association for Public Opinion Research (AAPOR), 2016. Standard Definitions: Final Dispositions of Case Codes and Outcome Rates for Surveys. Revised 2016. http://www.aapor.org/

Barbier, S., Wuyts, C., Italiano, P., & Loosveldt, G., 2016. European Social Survey Round 7 Belgium – Process evaluation for the data collection. Leuven: Centre for Sociological Research, KULeuven.

Bethlehem, J., Cobben, F., & Schouten, B., 2011. Handbook of Nonresponse in Household Surveys. Hoboken, New Jersey: John Wiley & Sons.

Blom, A., de Leeuw, E., Hox, J. Interviewer Effects on Nonresponse in the European Social Survey, 2010. MEA Discussion Paper No. 202-10. Available at SSRN: http://ssrn.com/abstract=1663704.

Campanelli, P., & O'Muircheartaigh, C., 1999. Interviewers, interviewer continuity, and panel survey nonresponse. Quality and Quantity, 33(1), 59–76.

Catania, J. A., Binson, D., Canchola, J., Pollack, L., Hauck, W., & Coates, T., 1996. Effects of interviewer gender, interviewer choice, and item wording on responses to questions concerning sexual behavior. Public Opinion Quarterly, 60, 345–375.

Chun, H., Tavarez, M., Dann, G., & Anastario, M., 2011. Interviewer gender and self-reported sexual behavior and mental health among male military personnel. International Journal of Public Health, 56(2), 225–229.

Davis, R., Couper, M., Janz, N., Caldwell, C., & Resnicow, K., 2010. Interviewer effects in public health surveys. Health Education Research, 25(1), 14–26.

Durrant, G., & D'Arrigo, J., 2014. Doorstep interactions and interviewer effects on the process leading to cooperation or refusal. Sociological Methods & Research, 43(3), 490–518.

Durrant, G., D'Arrigo, J. & Steele, F., 2011. Using field process data to predict best times of contact conditioning on household and interviewer influences. Journal of the Royal Statistical Society: Series A, 174, 1029–1049.

Durrant, G. B., Groves, R. M., Staetsky, L., & Steele, F., 2010. Effects of interviewer attitudes and behaviors on refusal in household surveys. Public Opinion Quarterly, 74(1), 1–36.

European Social Survey, 2012. Round 6 Data. Data file edition 2.1. Norwegian Social Science Data Services, Norway – Data Archive and distributor of ESS data for ESS ERIC.

European Social Survey, 2014. Round 7 Data. Data file edition 1.0. Norwegian Social Science Data Services, Norway – Data Archive and distributor of ESS data for ESS ERIC.

European Social Survey, 2016a. ESS-7 2014 Documentation Report. Edition 3.0. Bergen, European Social Survey Data Archive, NSD - Norwegian Centre for Research Data for ESS ERIC.

European Social Survey, 2016b. ESS-6 2012 Documentation Report. Edition 2.2. Bergen, European Social Survey Data Archive, NSD - Norwegian Centre for Research Data for ESS ERIC.

Fletcher, C., & Spencer, A., 1984. Sex of candidate and sex of interviewer as determinants of self-presentation orientation in interviews an experimental study. Applied Psychology, 33(3), 305–313.

Freeman, J., & Butler, E., 1976. Some sources of interviewer variance in surveys. Public Opinion Quarterly, 40(1), 79–91.

Groves, R., Cialdini, R., & Couper, M., 1992. Understanding the decision to participate in a survey. Public Opinion Quarterly, 56(4), 475–495.

Groves, R., & Couper, M., 1998. Nonresponse in household interview surveys. New York: Wiley.

Groves, R., Fowler Jr, F., Couper, M., Lepkowski, J., Singer, E., & Tourangeau, R. (2009). Survey methodology (Vol. 561). Hoboken, New Jersey: John Wiley & Sons.

Hansen, K., 2007. The effects of incentives, interview length, and interviewer characteristics on response rates in a CATI-study. International Journal of Public Opinion Research, 19(1), 112–121.

Hox, J., & de Leeuw, E., 2002. The influence of interviewers' attitude and behavior on household survey nonresponse: An international comparison. In R. Groves, D. Dillman. J. Eltinge, & R. Little (Eds.), Survey nonresponse (pp. 103–120). New York: Wiley.

Hox, J., de Leeuw, E., & Kreft, I., 1991. The effect of interviewer and respondent characteristics on the quality of survey data: a multilevel model. In P. Biemer, R. Groves, L. Lyberg, N. Mathiowetz, & S. Sudman. Measurement errors in surveys. (pp. 439-461) New York: Wiley.

Johnson, T., & Parsons, J., 1994. Interviewer effects on self-reported substance use among homeless persons. Addictive behaviors, 19(1), 83–93.

Katz, D., 1942. Do interviewers bias poll results?. Public Opinion Quarterly, 6(2), 248–268.

Lazarsfeld, P., & Merton, R. (1954). Friendship as a social process: A substantive and methodological analysis. Freedom and Control in Modern Society, 18(1), 18–66.

Lipman-Blumen, J., 1976. Toward a homosocial theory of sex roles: An explanation of the sex segregation of social institutions. Signs, 1(3), 15–31.

Lipps, O., 2010. Interviewer-respondent socio-demographic matching and survey cooperation. Survey Practice, 3(4).

Moorman, P., Newman, B., Millikan, R., Tse, C., & Sandler, D., 1999. Participation rates in a case-control study: The impact of age, race, and race of interviewer. Annals of Epidemiology, 9(3), 188–195.

O'Muircheartaigh, C., & Campanelli, P., 1998. The relative impact of interviewer effects and sample design effects on survey precision. Journal of the Royal Statistical Society: Series A (Statistics in Society), 161(1), 63–77.

Oyinlade, A., 2014. Models of likelihood of support for affirmative action. International Journal of Liberal Arts and Social Science, 2(5), 154–168.

Oyinlade, A., & Losen, A., 2014. Extraneous effects of race, gender, and race-gender homo- and heterophily conditions on data quality. SAGE Open, 4(1), 1–13

Pickery, J., & Loosveldt, G., 2002. A multilevel multinomial analysis of interviewer effects on various components of unit nonresponse. Quality and Quantity, 36(4), 427–437.

Pickery, J., Loosveldt, G., & Carton, A., 2001. The effects of interviewer and respondent characteristics on response behavior in panel surveys a multilevel approach. Sociological Methods & Research, 29(4), 509–523.

Pollner, M., 1998. The effects of interviewer gender in mental health interviews. The Journal of nervous and mental disease , 186, 369–73.

Rodgers, W., & Herzog, A., 1992. Collecting data about the oldest old: Problems and procedures. In Suzman, R., Willis, D., & Manton, K. (Eds.), The oldest old (pp. 135 – 156). Oxford University Press, USA.

Samples, T., Woods, A., Davis, T., Rhodes, M., Shahane, A., & Kaslow, N., 2014. Race of interviewer effect on disclosures of suicidal low-income African American women. Journal of Black Psychology, 40(1), 27–46.

Schnell, R., & Kreuter, F., 2005. Separating interviewer and sampling point effects. Journal of Official Statistics, 21(3), 389–410.

Singer, E., Frankel, M., & Glassman, M. (1983). The effect of interviewer characteristics and expectations on response. Public Opinion Quarterly, 47(1), 68–83.

Tirry, D., & Loosveldt, G., 2013. European Social Survey Round 6 Belgium Process evaluation for the data collection. Onderzoeksverslag Centrum voor Sociologisch Onderzoek (CeSO) Survey Methodology CeSO/ SM /2013-4.

Tourangeau, R., Rips, L. J., & Rasinski, K. A. (2000). *The psychology of survey response*. Cambridge, U.K.; New York: Cambridge University Press.

Tourangeau, R., & Yan, T., 2007. Sensitive questions in surveys. Psychological Bulletin, 133(5), 859–883.

van Tilburg, T., 1998. Interviewer effects in the measurement of personal network size. A nonexperimental study. Sociological Methods & Research, 26(3), 300–328.

Weeks, M., & Moore, R., 1981. Ethnicity-of-interviewer effects on ethnic respondents. Public Opinion Quarterly, 45(2), 245–249.

Williams, J. (1964). Interviewer-Respondent Interaction: A Study of Bias in the Information Interview. *Sociometry*, 27(3), 338-352.

Wilson, S., Brown, N., Mejia, C., & Lavori, P., 2002. Effects of interviewer characteristics on reported sexual behavior of California Latino couples. Hispanic Journal of Behavioral Sciences, 24(1), 38–62.

Table 1: Initial and final survey response outcome distribution and outcome rates in ESS6BE and ESS7BE

	ESS6BE				ESS7BE			
	Initial		Final		Initial		Final	
	N	%	N	%	N	%	N	%
Respondent (interview)	1,687	51.64	1,869	57.21	1,506	47.00	1,769	55.21
Refusal	902	27.61	778	23.81	965	30.12	837	26.12
Not able and other nonresponse	284	8.69	313	9.58	271	8.46	304	9.49
Noncontact	307	9.40	209	6.40	363	11.33	172	5.37
Ineligible	87	2.66	98	3.00	99	3.09	122	3.81
Total	3,267	100.0			3,204	100.0		
Response rate		53.05		58.98		48.50		57.40
Conditional refusal rate		31.40		26.28		35.19		28.76

Note: Initial survey response outcomes are based on the result of the initially assigned interviewer. The final survey response outcomes are based on the result at the end of the fieldwork. Response rates are calculated on the basis of AAPOR (2016) definition RR1, conditional refusal rates are calculated as the proportion of refusal sample units among eligible and contacted sample units.

Table 2: Top 10 item nonresponse in ESS6BE and ESS7BE

	ESS6BE			ESS7BE	
	Question label	**%**		**Question label**	**%**
1	Household's total net income, all sources	8.77	1	Father's highest level of education	10.68
2	Father's highest level of education	8.61	2	Household's total net income, all sources	8.48
3	Party voted for in last national election	7.37	3	Mother's highest level of education	6.73
4	Mother's highest level of education	6.63	4	Party voted for in last national election	4.96
5	Partner's highest level of education	4.07	5	Placement on left-right scale	4.07
6	Placement on left-right scale	3.32	6	Mother's occupation when respondent was 14	2.71
7	In Belgium, politicians take into account the views of other European governments	3.10	7	Father's occupation when respondent was 14	2.19
8	Mother's occupation when respondent was 14	2.60	8	Partner's highest level of education	2.01
9	In Belgium, citizens have the final say on political issues by voting directly in referendums	2.51	9	Trust in the United Nations	1.98
10	Importance of citizens having the final say on political issues by voting directly in referendums	2.41	10	Religious beliefs and practices are undermined or enriched by immigrants	1.64
	At least one of the shared 7 high-item nonresponse items	**26.06**		**At least one of the shared 7 high-item nonresponse items**	**25.66**

Note: Item nonresponse includes both Don't know and Refusal.

Table 3: Descriptive statistics sociodemographic characteristics of sample units (SU), respondents (Resp.) and interviewers (Int.)

	ESS6BE			ESS7BE		
	SU	**Resp.**	**Int.**	**SU**	**Resp.**	**Int.**
Male, %	48.18	48.75	51.30	49.11	50.65	51.33
Age, Mean (SD)	47.78 (19.39)	47.03 (19.10)	51.79 (11.92)	48.72 (19.21)	47.96 (18.95)	54.73 (12.04)
Education level, %						
Secondary not completed		31.07	10.39		30.85	12.00
Secondary		37.36	27.92		35.40	28.67
College		23.10	42.21		24.82	42.00
University		8.46	19.48		8.94	17.33
Job experience, %						
< 1 year			14.94			12.00
1 - 5 years			22.08			28.67
5 - 10 years			30.52			26.00
> 10 years			32.47			33.33
Missing, %	0.00	1.34	0.65	0.28	0.68	0.00
N	3,267	1,844	154	3,195	1,757	150

Note: The number of sample units and number of respondents included in the analysis differ slightly from the numbers reported in Table 1 due small numbers of missing values for some of the explanatory variables. Summations of percentages do not always add up to exactly 100% due to rounding.

Table 4: Sociodemographic matching between sample units (SU)/respondents (Resp.) and interviewers

	ESS6BE		ESS7BE	
	SU	**Resp.**	**SU**	**Resp.**
	%	**%**	**%**	**%**
Gender matching				
Male interviewer-male SU/respondent	24.48	24.95	24.63	26.47
Male interviewer-female SU/respondent	27.00	26.03	26.04	25.38
Female interviewer-male SU/respondent	23.69	23.81	24.48	24.19
Female interviewer-female SU/respondent	24.82	25.22	24.85	23.96
Age matching				
Age difference up to 5 years	17.88	18.00	16.24	18.27
SU/respondent older	33.82	34.11	30.86	31.02
Interviewer older	48.30	47.89	52.90	50.71
Education level matching				
Same level		27.28		26.29
SU/respondent higher		18.60		21.57
Interviewer higher		54.12		52.13
N	3,267	1,884	3,195	1,757

Note: Summations of percentages do not always add up to exactly 100% due to rounding.

Table 5: Multilevel logistic regression of sociodemographic matching on unit response and refusal

	ESS6BE		ESS7BE	
	Response	**Refusal**	**Response**	**Refusal**
	Est. (SE)	**Est. (SE)**	**Est. (SE)**	**Est. (SE)**
Gender matching				
Male interviewer-male sample unit	*(ref.)*	*(ref.)*	*(ref.)*	*(ref.)*
Male interviewer-female sample unit	-0.18 (0.10)	0.33** (0.12)	-0.16 (0.11)	0.10 (0.12)
Female interviewer-male sample unit	0.09 (0.14)	0.02 (0.14)	-0.05 (0.15)	-0.01 (0.14)
Female interviewer-female sample unit	0.14 (0.14)	-0.10 (0.14)	-0.01 (0.15)	0.04 (0.14)
Age matching				
Age difference up to 5 years	*(ref.)*	*(ref.)*	*(ref.)*	*(ref.)*
Sample unit older	0.07 (0.13)	-0.11 (0.14)	0.03 (0.13)	-0.25 (0.14)
Interviewer older	0.01 (0.12)	-0.20 (0.13)	-0.37** (0.13)	0.14 (0.14)
Sample unit: age (standardized)	-0.09 (0.07)	0.03 (0.07)	-0.24*** (0.07)	0.23** (0.07)
Interviewer: education level				
Secondary not completed	*(ref.)*	*(ref.)*	*(ref.)*	*(ref.)*
Secondary	-0.27 (0.21)	0.06 (0.20)	0.31 (0.21)	-0.32 (0.19)
College	-0.24 (0.20)	-0.04 (0.19)	0.20 (0.20)	-0.23 (0.17)
University	-0.26 (0.23)	-0.24 (0.22)	-0.05 (0.23)	0.09 (0.20)
Interviewer: job experience				
< 1 year	*(ref.)*	*(ref.)*	*(ref.)*	*(ref.)*
1 - 5 years	-0.35 (0.19)	0.21 (0.18)	0.34 (0.22)	-0.24 (0.19)
5 - 10 years	-0.29 (0.18)	0.14 (0.18)	0.52* (0.22)	-0.40* (0.20)
> 10 years	-0.36* (0.18)	0.24 (0.18)	0.35 (0.22)	-0.31 (0.19)
Intercept variance	0.26	0.17	0.31	0.16
Log likelihood ratio test versus Model0	$\chi^2(5) = 6.79$, $p = 0.24$	$\chi^2(5) = 15.24$, $p < 0.01$	$\chi^2(5) = 11.70$, $p = 0.04$	$\chi^2(5) = 6.69$, $p = 0.24$
N	3,162	2,855	3,096	2,736

*Note: Coefficients reported with standard errors in parentheses. The corresponding Model0 that excludes the matching variables is used for model comparison. * p < 0.05, ** p < 0.01, *** p < .001.*
N refers to the number of eligible sample units (Response) or the number of contacted eligible sample units (Refusal) for which complete information on explanatory variables is available.

Table 6: Multilevel logistic regression of sociodemographic matching on item nonresponse to at least one of the 6 common item nonresponse items

	ESS6BE	ESS7BE
	Est. (SE)	**Est. (SE)**
Gender matching		
Male interviewer-male respondent	*(ref.)*	*(ref.)*
Male interviewer-female respondent	0.42* (0.17)	0.45** (0.17)
Female interviewer-male respondent	0.26 (0.20)	0.48** (0.18)
Female interviewer-female respondent	0.84*** (0.21)	0.43* (0.19)
Age matching		
Age difference up to 5 years	*(ref.)*	*(ref.)*
Respondent older	0.54** (0.20)	0.72*** (0.20)
Interviewer older	0.48* (0.20)	-0.19 (0.20)
Education level matching		
Same level	*(ref.)*	*(ref.)*
Respondent higher	-0.07 (0.22)	-0.07 (0.21)
Interviewer higher	0.05 (0.17)	0.12 (0.17)
Respondent: age (standardized)	-0.06 (0.09)	-0.58*** (0.10)
Respondent: education level		
Secondary not completed	*(ref.)*	*(ref.)*
Secondary	-0.29 (0.15)	-0.55*** (0.15)
College	-0.60** (0.21)	-0.96*** (0.21)
University	-0.76* (0.33)	-1.26*** (0.34)
Interviewer: job experience		
< 1 year	*(ref.)*	*(ref.)*
1 - 5 years	0.16 (0.27)	-0.29 (0.27)
5 - 10 years	0.13 (0.26)	-0.24 (0.27)
> 10 years	-0.06 (0.26)	-0.16 (0.26)
Intercept variance	0.49	0.24
Log likelihood ratio test versus Model0	$\chi^2(7) = 34.47$, $p < 0.001$	$\chi^2(7) = 29.38$, $p < 0.001$
N	1,844	1,750

*Note: Coefficient estimates reported with standard errors in parentheses. The corresponding Model0 that excludes the matching variables is used for model comparison. * p < 0.05, ** p < 0.01, *** p < .001*